

# Artificial intelligence for the prediction of water quality index in groundwater systems

Mohamad Sakizadeh<sup>1</sup>

Received: 6 October 2015 / Accepted: 8 December 2015 / Published online: 22 December 2015  
© Springer International Publishing Switzerland 2015

**Abstract** A study was initiated to predict water quality index (WQI) using artificial neural networks (ANNs) with respect to the concentrations of 16 groundwater quality variables collected from 47 wells and springs in Andimeshk during 2006–2013 by the Iran's Ministry of Energy. Such a prediction has the potential to reduce the computation time and effort and the possibility of error in the calculations. For this purpose, three ANN's algorithms including ANNs with early stopping, Ensemble of ANNs and ANNs with Bayesian regularization were utilized. The application of these algorithms for this purpose is the first study in its type in Iran. Comparison among the performance of different methods for WQI prediction shows that the minimum generalization ability has been obtained for the Bayesian regularization method (MSE = 7.71) and Ensemble averaging method (MSE = 9.25), respectively and these methods showed the minimum over-fitting problem compared with that of early stopping method. The correlation coefficients between the predicted and observed values of WQI were 0.94 and 0.77 for the test and training data sets, respectively indicating the successful prediction of WQI by ANNs through Bayesian regularization algorithm. A sensitivity analysis was implemented to show the importance of each parameter in the prediction of WQI during ANN's modeling and showed that parameters like Phosphate and Fe are the most influential parameters in the prediction of WQI.

**Keywords** Water quality index · Artificial neural network · Early stopping · Ensemble method · Bayesian regularization

## Introduction

Since the analysis of environmental data is inherently complex, with data sets often containing nonlinearities; temporal, spatial, and seasonal trends; and non-gaussian distributions so, neural networks as nonlinear mapping structures can be used in this field. For instance, neural networks provide a powerful inference engine for regression analysis, which stems from the ability of neural networks to map nonlinear relationships, that is more difficult and less successful when using conventional time-series analysis (May et al. 2009). As a result of high capability of artificial neural networks (ANNs), numerous applications of predictive neural network models to environmental studies have been reported in the literature. Interested readers can refer to Hanrahan (2011) for a detail list of published works in different aspects of environmental sciences. There is also a comprehensive application of ANNs in water quality researches. For instance Singh et al. (2009) studied the capability of ANNs for computing the dissolved oxygen (DO) and biochemical oxygen demand (BOD) levels in the Gomti River (India). In this study two ANN models using the coefficient of determination ( $R^2$ ; square of the correlation coefficient), root mean square error (RMSE) and bias as the performance criteria were considered. Both of the models employed eleven input water quality variables measured in river water over a period of 10 years each month at eight different sites. The model computed values of DO and BOD by both the ANN models were in close agreement with their respective

✉ Mohamad Sakizadeh  
msakizadeh@gmail.com

<sup>1</sup> Department of Environmental Sciences, Faculty of Sciences, Shahid Rajaei Teacher Training University, Tehran, Iran

measured values in the river water. On the contrary, Dogan et al. (2009) investigated the abilities of an artificial neural networks' (ANNs) model to improve the accuracy of the biological oxygen demand (BOD) estimation in Melen River in Turkey. The prediction was implemented using 11 water quality variables (chemical oxygen demand, temperature, dissolved oxygen, water flow, chlorophyll a and nutrients, ammonia, nitrite, nitrate) that affect biological oxygen demand concentrations at 11 sampling sites during 2001–2002. Different hidden nodes were tried to find the best neural network structure and finally the ANN architecture having eight inputs and one hidden layer with three nodes gave the best choice. Comparison of the results revealed that the ANN model gives reasonable estimates for the BOD prediction. Khalil et al. (2011) used three models including ANN, ensemble of ANNs and canonical correlation analysis for the estimation of water quality mean values at un-gauged sites. The outputs of these models consisted of four water quality variables. The three models were applied to 50 sub-catchments in the Nile Delta, Egypt. A jackknife validation procedure was used to evaluate the performance of the three models. The results showed that the ensemble of ANN model provides better generalization ability than the ANN.

A numbers of indices have been developed all over the world to consider the overall quality of water given a number of variables (Ayorlo and Ramdzani 2014). These indices are based on the comparison of the water quality parameters to regulatory standards and give a single value to the water quality of a source (Khan et al. 2003). The water quality index (WQI) may be defined as a single numeric score that describes the water quality condition at a particular location in a specific time (Kaurish and Younos 2007). These indices offer several advantages including representation of measurements on many variables varying in measurement units in one metric, thus establishing a criterion for tracking changes in water quality over time and space and simplifying communication of the monitoring results (Gazzaz et al. 2012). Water quality index is one of the most effective tools to monitor the surface as well as ground water pollution and can be used efficiently in the implementation of water quality upgrading programmes. The objective of an index is to turn multifaceted water quality data into simple information that is comprehensible and useable by the public. It is one of the aggregate indices that have been accepted as a rating that reflects the composite influence on the overall quality of numbers of precise water quality characteristics (Tiwari and Mishra 1985). Water quality index provides information on a rating scale from zero to hundred. Higher value of WQI indicates better quality of water and lower value shows poor water quality (Rupal et al. 2012). The prediction of WQI using artificial neural networks has been

implemented earlier in different countries (Khuan et al. 2002; Rene and Saidutta 2008; Gazzaz et al. 2012). This study was initiated to predict the performance of ANNs using three algorithms with different neural structures. Such a prediction has the potential to reduce the computation time and effort and the possibility of errors in the calculation. The application of these algorithms for this purpose is the first study in its type in Iran.

## Materials and methods

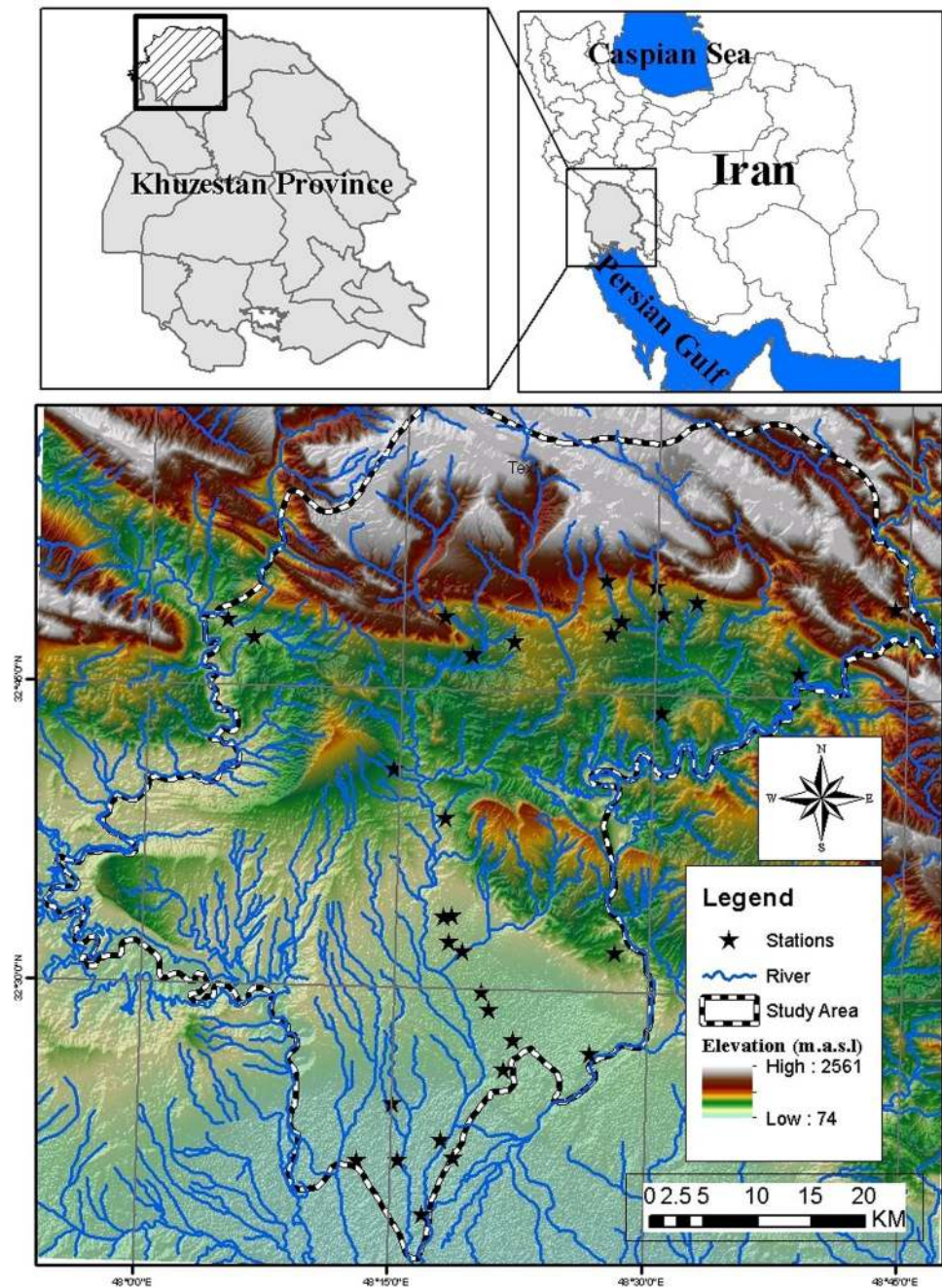
### Study area

Andimeshk is a city in the northern part of Khuzestan Province on the southern slopes of Zagross Mountains in the southwest of Iran. The total study area is about 3100 km<sup>2</sup>. According to the latest census by Iranian Statistical Center in 2012, the total population of this city is 167,126 inhabitants among them 128,774 inhabitants reside in urban area and 34,985 inhabitants live in rural area, respectively. The average annual precipitation in the region is about 353 mm and in 2012, the average exploitation from groundwater resources was 133,981 thousand cubic meters/year according to the report prepared by Khuzestan Water and Power Authority. The major irrigated broad-acre crops grown in this area are wheat, barley and maize, in addition to fruits, melons, watermelons and vegetables such as tomatoes and cucumbers (Rezania et al. 2009). Average water-level fluctuations in this region is very low, about 0.5–1 m between dry and wet seasons because of continuous recharge with Dez and Karkhe Rivers (Nouri et al. 2006). A view of the study area has been illustrated in Fig. 1.

### Water quality index calculation

In this study, we used 16 groundwater quality variables [EC, TDS, Turbidity, pH, Total hardness, Ca, Mg, Sulfate, Phosphate, Nitrate, Nitrite, Fluoride, Fe, Mn, Cu, Cr(VI)] in 47 wells and springs during an 8 years time period (2006–2013) gathered by the Iran Ministry of Energy to calculate the water quality index followed by ANN modeling of this index in Andimeshk located in southwest of Iran. To calculate the water quality index, the method proposed by Horton (1965), and followed by many researchers (e.g. Dwivedi and Pathak 2007; Patel and Desai 2006) was used. Concisely, in this method, rating scale (in the form of step functions) is assigned to each water quality parameter to be considered and each parameter is weighted according to its relative significance in overall quality. The following formula is used to calculate the water quality index (WQI):

**Fig. 1** An illustration of study area including Khuzestan Province and sampling stations



$$WQI = \sum q_i w_i / \sum w_i \tag{1}$$

In which  $q_i$  is the rating scale and  $w_i$  is the related unit weight for the respective water quality parameter.

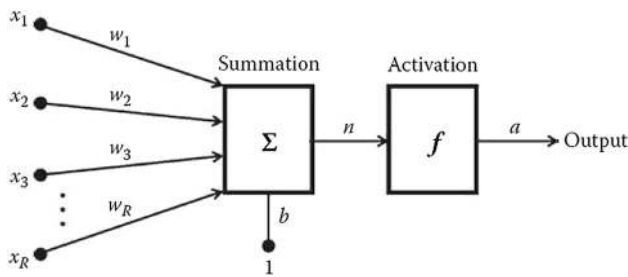
**Artificial neural network modeling**

The output of a neural network relies on the functional cooperation of the individual neurons within the network, where processing of information is characteristically done in parallel rather than sequentially as in earlier binary

computers (Hanrahan 2010). The building blocks of a neural network consists of individual scalar inputs  $x_1, x_2, x_3, \dots, x_n$  which are each weighted with appropriate elements  $w_1, w_2, w_3, \dots, w_n$  of the weight matrix  $W$ . The sum of the weighted inputs and the bias forms the net input  $n$ , proceeds into a transfer function  $f$ , and produces the scalar neuron output  $a$  written as (Hagan et al. 1996):

$$a = f(WX + b) \tag{2}$$

A view of these operations excerpted from Hanrahan (2011) can be found in Fig. 2. Linear and nonlinear transfer



**Fig. 2** An illustration of a typical artificial neural network structure

functions are usually used for this purpose in which the choice of transfer function strongly influences the complexity and performance of neural networks (Hagan et al. 1996). The most well known transfer functions are Linear, Log-sigmoid and Hyperbolic tan-sigmoid transfer functions. In linear transfer function, the function computes a neuron’s output by merely returning the value passed directly to it. Log-sigmoid transfer functions take given inputs and generate outputs between 0 and 1 as the neuron’s net input goes from negative to positive infinity which mathematically expressed as:

$$a = \frac{1}{1 + e^{-n}} \tag{3}$$

This transfer function is usually utilized in back-propagation networks. Gaussian-type (hyperbolic tan-sigmoid transfer function) functions are employed in radial basis function networks and are mathematically expressed as:

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}} \tag{4}$$

In the above equations  $n$  is the net input to the neuron. Three neural network algorithms were utilized for the prediction of WQI given sixteen groundwater quality variables. (a) artificial neural network with early stopping (b) ensemble of artificial neural networks (c) artificial neural network with Bayesian regularization.

*Artificial neural network with early stopping*

To keep within the scope of this paper, we limited our survey of NN models to the feed-forward neural network with one hidden node. The linear transfer function and the following transfer function were used for the output and hidden layers respectively:

$$y_j = \tanh \left( \sum_{i=1}^d w_{ij}x_i + b_j \right) \tag{5}$$

where  $w_{ij}$  and  $b_j$  are the weight and bias parameters in which “ $i$ ” and “ $j$ ” subscripts refer to the input and neuron, respectively. In addition, Levenberg–Marquardt algorithm

was used to update the weight and bias of the network according to this formula:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{6}$$

where  $J$  is the Jacobian matrix, which contains first derivatives of the network errors with respect to the weights and biases,  $e$ , is a vector of network errors,  $I$  is the identity matrix,  $x$ , is a vector containing weights and biases, and  $\mu$  is a scalar value, respectively. The way that available data are divided into training, testing, and validation subsets can have a significant influence on the performance of an artificial neural network. Despite numerous studies, no systematic approach has been developed for the optimal division of data for ANN models (Bowden et al. 2002). So, in order to avoid over-fitting of the model to the training data, different data division was tried and the original data were separated into training and test set. In addition, since each backpropagation training session starts with different initial weights and biases therefore, each neural network was trained 20 times starting from different initial weight and biases and the average performance of the networks based on mean squared error(MSE) was considered. In order to avoid over-fitting of the model to the training data *early stopping* was used.

In this technique the available data are divided into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over-fit the data, the error on the validation set typically begins to rise. When the validation error increases for a specified number of iterations the training is stopped, and the weights and biases at the minimum of the validation error are returned (Beale et al. 2013). Since the water quality variables had different scales which makes problem for neural network modeling so, they were standardized before modeling. One of the parameters that affect the generalization ability of a neural network is the number of hidden nodes which is usually optimized through a trial and error procedure (Coulibaly et al. 2000), thus the same procedure was followed to obtain the optimum hidden nodes in this study.

*Artificial neural network with ensemble averaging*

One of the methods to improve generalization ability of neural network especially for noisy data or small data set like that of this study is to train multiple neural networks and average their outputs. The ANN ensemble is a group of ANNs that are trained for the same problem, where results

obtained by these ANNs are combined to produce the ANN ensemble output (Khalil et al. 2011). Several methods were proposed in the literature to generate ensemble ANNs: (1) manipulating the set of initial random weights, (2) using different network topology, (3) training component networks using different training algorithms, and (4) manipulating the training set (Shu and Ouarda 2007). In this study, the first method was utilized and the out-of-sample performance of the ensemble deemed to be lower than most of the individual performances.

*Artificial neural network with Bayesian regularization*

This approach minimizes the over-fitting problem by taking into account the goodness-of-fit as well as the network architecture. The Bayesian regularization approach involves modifying the usually used objective function, such as the mean sum of squared network errors (MSE):

$$E_d = \sum_{i=1}^n (t_i - a_i)^2 \tag{7}$$

Here n is the number of inputs in the training set, t<sub>i</sub> is the i<sup>th</sup> expected output and a<sub>i</sub> is the i<sup>th</sup> output obtained as neural network response.

The modification aims to improve the model’s generalization capability. The objective function in the above equation is expanded with the addition of a term, E<sub>w</sub> which is the sum of squares of the network weights:

$$F = \beta E_d + \alpha E_w \tag{8}$$

where E<sub>w</sub> is the sum of squares of the network weights and α and β are objective function parameters. The smaller the weights the better is the generalization capability of the network (Kumar et al. 2004). In this approach, the weights and biases of the network are assumed to be random variables with specified distributions (Daliakopoulos et al. 2005). All of the required programs in this study have been written in MATLAB (R2013b).

**Results and discussion**

The results of ANN modeling of WQI for different number of hidden nodes and different data divisions and methods have been rendered in Table 1. Considering this table, the best data division is 80 % for the training data which has resulted in the minimum training and generalization error among all of the methods. Since the ANNs are unable to extrapolate beyond the range of the data used for training (Tokar and Johnson 1992) so, for the best generalization (out-of-sample) ability, the training and test set must be representative of the same population (Masters 1995) therefore the previous researches have proved that the way that the available data are divided into subsets can have a significant influence on an ANN’s performance (e.g. Maier and Dandy 1996; Tokar and Johnson 1992). In this respect, the method of data division was studied by Maier and

**Table 1** The results of water quality index prediction with artificial neural network using three different algorithms

% of training data	Number of hidden nodes	Early stopping		Ensemble averaging		Bayesian regularization	
		Training MSE	Testing MSE	Training MSE	Testing MSE	Training MSE	Testing MSE
60	5	17.04	86.60	10.13	79.18	0.01	27.90
	10	20.18	103.13	7.78	84.66	0.23	35.30
	15	5.52	82.26	1.68	72.71	8.63	15.73
	20	21.15	40.90	3.55	15.65	5.53	14.86
	25	14.29	132.47	4.06	88.13	5.86	12.29
	30	41.57	41.48	7.53	19.50	10.74	42.30
70	5	24.69	41.52	10.46	81.51	1.44	82.47
	10	20.65	126.03	8.37	103.14	2.23	19.15
	15	39.70	36.69	12.16	14.20	7.87	11.75
	20	29.83	66.60	9.70	33.31	8.74	13.36
	25	34.48	58.41	7.10	21.27	8.86	8.93
	30	23.94	97.13	5.41	70.91	8.38	44.91
80	5	36.72	68.62	12.41	27.67	0.17	24.93
	10	24.47	48.16	6.40	21.70	0.38	26.52
	15	55.32	32.20	24.52	6.89	0.79	31.80
	20	21.31	17.67	3.52	9.25	8.33	13.88
	25	23.45	145.07	4.65	10.75	8.08	7.71
	30	11.88	114.40	3.53	96.56	6.70	27.15

Dandy (2000) through considering 43 papers in which ANNs have been used for the prediction of water resources variables. It was concluded that in only two out of 43 papers, an independent test set was used in addition to training and validation sets and for the rest of the case studies data division was performed incorrectly. That is way in the current research, the data set were initially divided into training and test set to ensure that a completely independent part is being used as the test set.

The results of this study also confirm the importance of data division on ANN's prediction ability. In this field, Ray and Klindworth (2000) believe that in order to achieve the best generalization ability in ANN's modeling, the training data must be large enough to represent the full population and the testing set should be randomly chosen and manually checked to ensure that they had characteristics similar to the training data. This procedure has also been confirmed in this study as the higher percent of the training set has resulted in better performance and the testing set has been chosen randomly to be representative of the original training set.

Comparison among the performance of different methods for WQI prediction shows that the minimum generalization ability has been obtained for the Bayesian regularization method ( $MSE = 7.71$ ) and Ensemble averaging method ( $MSE = 9.25$ ), respectively. The minimum generalization ability of Early stopping method was significantly higher than the two above-mentioned methods ( $MSE = 17.67$ ) anyhow. The minimum training MSE was 0.01 and was related to Bayesian regularization method however, the respective testing MSE was significantly higher ( $MSE = 27.90$ ) indicating the over-fitting of the model to the training data set. Over-fitting problem or poor generalization capability happens when a neural network over learns during the training period. As a result, such a too well-trained model may not perform well on unseen data set due to its lack of generalization capability (Doan and Liong 2004).

Some of the previous researches have proved that the use of appropriately selected ensemble techniques allows to improve the performance of a single ANN (e.g. Ouarda and Shu 2009; Zaier et al. 2010). One of redeeming features of using an ensemble of ANNs is that it avoids the local minima problem by averaging in functional space not parameter space (Perrone and Cooper 1992). An additional benefit of the ensemble method's ability to combine multiple regression estimates is that the regression estimates can come from many different sources. This fact allows for great flexibility in the application of the ensemble method. For example, the networks can have different architectures or be trained by different training algorithms or be trained on different data sets (Perrone and Cooper 1992). Prediction of water quality index using artificial intelligence has

been implemented in earlier studies. For instance, Lermontov et al. (2009) proposed the creation of a new water quality index based on fuzzy logic, called fuzzy water quality index (FWQI). This index showed a good correlation with the WQI traditionally calculated in Brazil. This study showed that fuzzy logic can be used to develop a meaningful water quality index, since the FWQI is comparable to other indices proposed in the literature. Application of fuzzy set theory to compute the water quality index has also been confirmed in the study of Nasiri et al. (2007).

One of the problems of using small data set like that of this paper is the danger of over-fitting which is exacerbated when using leave-one-out method meaning that a small part of data are retained for the testing set. The problem is that since the network is never trained on the hold-out data the network may be missing valuable information about the distribution of the data. This problem is aggravated when limited data are available as it might be difficult to assemble a representative validation set because the training and validation sets should be representative of the same population (Maier and Dandy 2000).

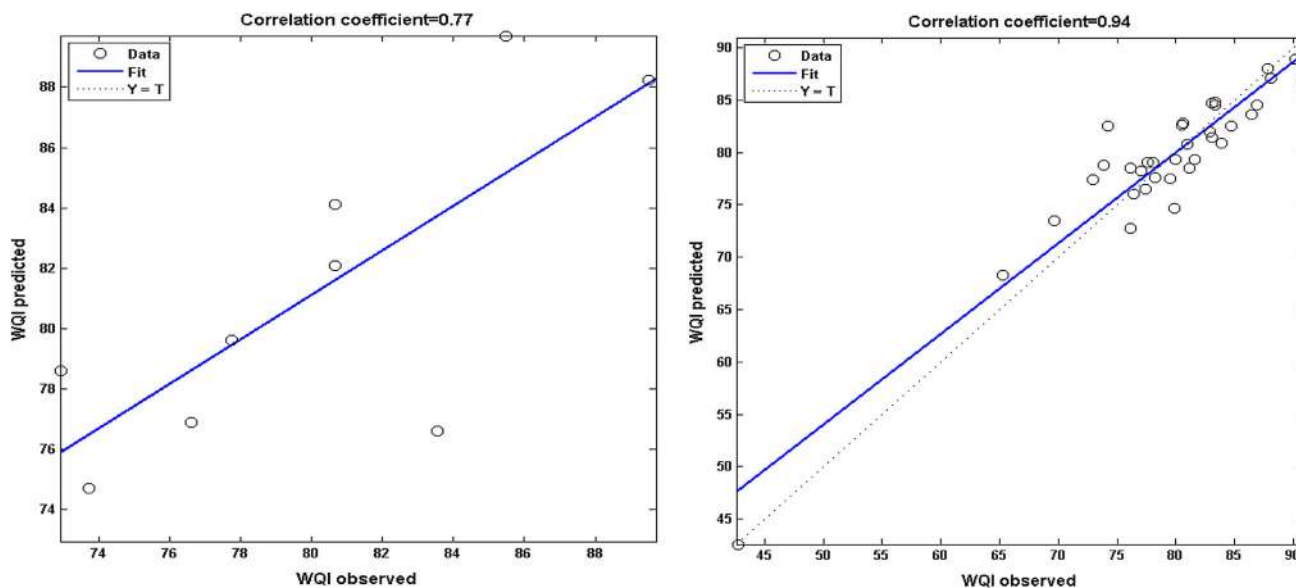
This is not a problem with ensemble method because we have used different data division for different ANNs used in the ensemble. In this way, the population as a whole will use the entire data set while each network has avoided over-fitting by using a cross-validators stopping rule. The other benefit of using ensemble of neural network is the solution of the problem of local minima which is a common problem when modeling with single neural networks. It means that during the neural network training with a single ANNs, the model gets stuck into local instead of global minima. A simple way to deal with the local minima problem is to use an ensemble of ANNs, where their parameters are randomly initialized before training. The individual NN solutions would be scattered around the global minimum, but by averaging the solutions from the ensemble members, we would likely obtain a better estimate of the true solution (Hsieh and Tang 1998). Although ensemble methods are useful in addressing the problem of over-fitting however there are some other techniques which have been used in the literature. For example Sakizadeh et al.(2015) have used bootstrap and noise injection methods to address the problem of small data record and over-fitting in the prediction of water quality index using ANN methods. On the contrary, the problem of over-fitting in short records of noisy data was obviated by adding a complexity term to the cost function besides using validation methods in the study of Weigend et al. (1990).

The results of the previous studies on generalization ability of Bayesian regularization methods in comparison with that of early stopping have shown that the Bayesian regularization approach lends the model higher

generalization ability (Doan and Liong 2004). In addition, the prediction improvement of models trained with generalization approaches over those with standard approach is considerably remarkable when the data is very noisy (Doan and Liong 2004). The high performance of this method confirms that of the earlier studies. One of the features of ANNs modeling using Bayesian regularization is that it controls over-fitting even with a large hidden layer (Anctil et al. 2004). As a whole, the number of hidden neurons has a direct impact on the prediction capability of modeling with ANNs. Setting too few hidden units causes high training errors and high generalization errors due to under-

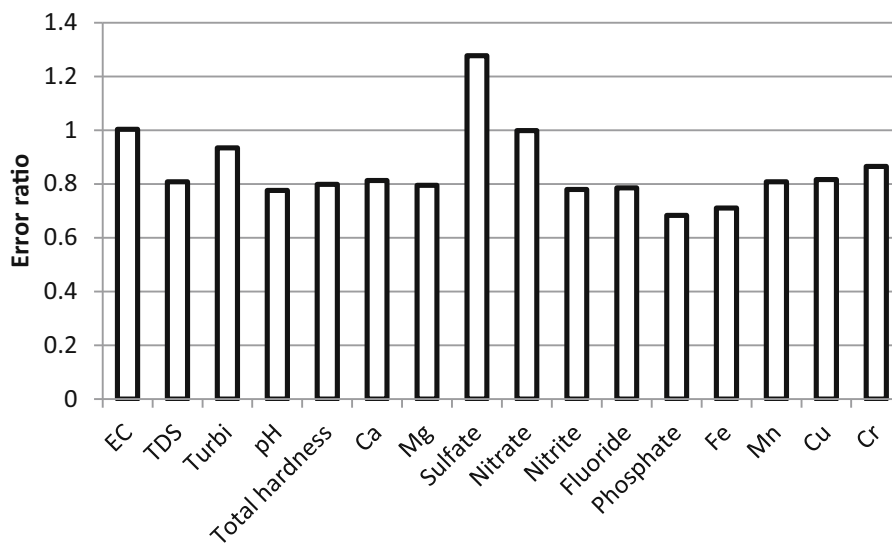
fitting, while too many hidden units result in low training errors but still high generalization errors due to over-fitting (Xu and Chen 2008).

Having done the modeling of WQI by ANNs, in the next step, given the optimum network structure and the best algorithm, the predicted WQI was compared with that of the observed value for the training and testing set (Fig. 3). According to this Figure, the correlation coefficients between the predicted and observed values of WQI were 0.94 and 0.77 for the test and training data sets, respectively indicating the successful prediction of WQI by ANN through Bayesian regularization algorithm.



**Fig. 3** Comparison between the WQI observed and predicted by ANNs with Bayesian regularization for the test (*left*) and training (*right*) data sets

**Fig. 4** The results of sensitivity analysis for groundwater quality variables



In order to determine the most important parameters in the prediction of WQI, a sensitivity analysis was implemented using the error ratio between the original MSE and the MSE after exclusion of each groundwater quality parameter from the list of parameters (Fig. 4). With respect to this Figure, it can be concluded that the least important parameters are Sulfate, EC and Nitrate, respectively whereas some other parameters like Phosphate and Fe are the most influential parameters in the prediction of WQI. The significance of the other parameters was intermediate and roughly in the same range. The most important parameters are those for which the error ratio was higher when they had been excluded from the model indicating the importance of these parameters in modeling by ANN while the least important parameters only exerted a minor error when removed from the model. This analysis is helpful for the identification of less important variables to be removed or ignored in subsequent studies in addition to the most essential variables (Gazzaz et al. 2012).

## Conclusion

One of the problems of ANN's modeling in environmental studies which suffer from the problem of small data record is the danger of over-fitting of the model to the training data resulting in poor generalization of the model for the data out-of-the training data range. The results of this study proved that using some algorithms like Bayesian regularization and Ensemble methods this problem can be obviated. The prediction of water quality index (WQI) was successfully implemented by Bayesian regularization and Ensemble averaging methods, though the performance of Bayesian regularization was roughly better, with minimum test error indicating the good generalization ability of these methods in this field. The poor generalization ability is a problem which has been overlooked by most of the researches in all around of the world although it is an important issue which should be taken into account.

**Acknowledgments** The authors are grateful to the help of Andimeshk Health Network and Iran Ministry of Energy for providing us with water quality data.

## References

- Ajorlo M, Ramdzani A (2014) Assessment of stream water quality in tropical grassland using water quality index. *Ecopersia* 2:427–440
- Anctil F, Perrin C, Andreassian V (2004) Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ Model Soft* 19:357–368
- Beale MH, Hagan MT, Demuth HB (2013) Neural network toolbox user guide. The MathWorks, Inc, USA
- Bowden GJ, Maier HR, Dandy GC (2002) Optimal division of data for neural network models in water resources applications. *Water Resour Res*. doi:10.1029/2001WR000266
- Coulibaly P, Anctil F, Bobee B (2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J Hydrol* 230:244–257
- Daliakopoulos IN, Coulibaly P, Tsanis IK (2005) Groundwater level forecasting using artificial neural networks. *J Hydrol* 309:229–240
- Doan CD, Liong SY (2004) Generalization for multilayer neural network Bayesian regularization or early stopping. The 2nd APHW Conference, Singapore
- Dogan E, Sengorur B, Koklu R (2009) Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. *J Environ Manage* 90:1229–1235
- Dwivedi SL, Pathak V (2007) A preliminary assignment of water quality index to Mandakini river, Chitrakoot. *Ind. J Environ Prot* 27:1036–1038
- Gazzaz NM, Yusoff MK, Aris AZ, Juahir H, Ramli MF (2012) Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar Pollut Bull* 64:2409–2420
- Hagan MT, Demuth HB, Beale MH (1996) Neural network design. PWS Publishing Company, Boston
- Hanrahan G (2010) Computational neural networks driving complex analytical problem solving. *Anal Chem* 82:4307–4313
- Hanrahan G (2011) Artificial neural network in biological and environmental analysis. Taylor and Francis Group, USA
- Horton RK (1965) An index number system for rating water quality. *J Water Pollut Control Fed* 37:300–305
- Hsieh WW, Tang B (1998) Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull Am Meteorol Soc* 79:1855–1870
- Kaurish FW, Younos T (2007) Developing a standardized water quality index for evaluating surface water quality. *J Am Water Resour Assoc* 43:533–545
- Khalil B, Ouarda TBMJ, St-Hilaire A (2011) Estimation of water quality characteristics at un-gauged sites using artificial neural networks and canonical correlation analysis. *J Hydrol* 405:277–287
- Khan F, Husain T, Lumb A (2003) Water quality evaluation and trend analysis in selected watersheds of the Atlantic region of Canada. *Environ Monit Assess* 88:221–242
- Khuan LY, Hamzeh N, Jailani R (2002) Prediction of water quality index (WQI) based on artificial neural network (ANN), student conference on research and development proceedings, Shah Alam, Malaysia
- Kumar P, Merchant SN, Desai UB (2004) Improving performance in pulse radar detection using Bayesian regularization for neural network training. *Digit Signal Process* 14:438–448
- Lermontov A, Yokoyama L, Lermontov M, Machado MAS (2009) River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed, Brazil. *Ecol Indic* 9(6):1188–1197
- Maier HR, Dandy GC (1996) The use of artificial neural networks for the prediction of water quality parameters. *Water Resour Res* 32:1013–1022
- Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environ Modell Softw* 15:101–124
- Masters T (1995) Advanced algorithms for neural networks. New York, Wiley, USA
- May RJ, Maier HR, Dandy GC (2009) Developing artificial neural networks for water quality modelling and analysis. In: Hanrahan



- G (ed) Modelling of pollutants in complex environmental systems. ILM Publications, St. Albans
- Nasiri F, Maqsood I, Huang G, Fuller N (2007) Water quality index: a fuzzy river-pollution decision support expert system. *J Water Resour Plann Manage* 133(2):95–105
- Nouri J, Mahvi AH, Babaei AA, Jahed GR, Ahmadpour E (2006) Investigation of heavy metals in groundwater. *Pak J Biol Sci* 9:377–384
- Ouarda TBMJ, Shu C (2009) Regional low-flow frequency analysis using single and ensemble artificial neural networks. *Water Resour Res* 45:W11428
- Patel S, Desai KK (2006) Studies on water quality index of some villages of Surat district. *Pollut Res* 25:281–285
- Perrone MP, Cooper LN (1992) When networks disagree: ensemble methods for hybrid neural networks. Technical report, Institute for Brain and Neural System & Brown University Providence, Rhode Island
- Ray C, Klindworth KK (2000) Neural networks for agrichemical vulnerability assessment of rural private wells. *J Hydrol Eng* 5:162–171
- Rene ER, Saidutta MB (2008) Prediction of water quality indices by regression analysis and artificial neural networks. *Int J Environ Res* 2(2):183–188
- Rezania AR, Naseri AA, Albaji M (2009) Assessment of soil properties for irrigation methods in North Andimeshk Plain, Iran. *J Food Agric Environ* 7:728–733
- Rupal M, Tanushree B, Sukalyan C (2012) Quality characterization of groundwater using water quality index in Surat city, Gujarat, India. *Int Res J Environ Sci* 1:14–23
- Sakizadeh M, Malian A, Ahmadpour E (2015) Groundwater quality modeling with a small data set. *Ground Water*. doi:10.1111/gwat.12317
- Shu C, Ouarda TBMJ (2007) Flood frequency analysis at un-gauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resour Res* 43:W07438
- Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality-A case study. *Ecol Model* 220:888–895
- Tiwari TN, Mishra MA (1985) A preliminary assignment of water quality index of major Indian rivers. *Indian J Environ Prot* 5:276–279
- Tokar AS, Johnson PA (1992) Rainfall-runoff modeling using artificial neural networks. *J Hydrol Eng* 4:232–239
- Weigend AS, Huberman BA, Rumelhart DE (1990) Predicting the future: a connectionist approach. *Int J Neur Syst* 1(3):193–209
- Xu S, Chen L (2008) A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. 5th international conference on information technology and applications, Australia
- Zaier I, Shu C, Ouarda TBMJ, Seidou O, Chebana F (2010) Estimation of ice thickness on lakes using artificial neural networks ensembles. *J Hydrol* 380:330–340